**Narendra Malkar**

**Email: narendramalkar2@gmail.com**

**Mobile No: +91- 9029845812**

**Address: Pune, MH.**

***SUMMARY***

Experienced Data Engineer and Analytics with **7 years** in areas of development, data Pipeline, data Lake, data warehousing and data analytics. Proven track record in leveraging advanced technologies to enhance data insights and optimize performance.

***PROFESSIONAL SUMMARY***

* **Data Architecture and Implementation:** Architected and managed scalable data lake solutions using different storage account services. Organizing data into raw, processed, and curated layers(**Medallion Architecture**). This resulted in a 30% improvement in retrieval times and a 20% reduction in storage costs.
* **Data Pipeline and ETL:** Designed, developed, and managed end-to-end data pipelines using Aws Glue to automate data movement and transformation from Data Lake Storage (S3) to PySpark and finally to RedShift for analytics and reporting.
* **Data Validation and Cleansing:** Implemented robust data validation and rule-based cleansing techniques using PySpark to standardize data formats, ensuring accuracy and consistency throughout the data pipeline.
* Developed data quality frameworks in PySpark to detect, log, and resolve anomalies (such as missing values, duplicates, and inconsistencies), leading to a 25% improvement in data accuracy.
* **Data Modeling and Optimization:** Developed & optimized data models (star, snowflake) and implemented data warehouse to support business intelligence and analytics requirements.

Enhanced query performance & scalability through **indexing**, **partitioning** & **views**, reducing times by 40%.

* **Data transformation and Aggregation:** Implemented transformation logic to convert raw data into structured, meaningful datasets that align with business requirements. Developed data transformation and aggregation processes using SQL & PySpark scripts, improving processing efficiency by 30% compared to ETL methods.
* **Optimization and Performance Tuning:** Implemented PySpark optimizations including partitioning, parallel processing, and broadcast joins to enhance ETL performance, reducing processing times by 50%. Utilized monitoring and Log Analytics for proactive performance management.
* **End-to-End Analytics Solution:** Developed analytics solutions by integrating Data Lake, DWH with data Analytics techniques, facilitating advanced reporting and BI dashboards in Power BI, improving decision-making processes.
* **Project Highlights:** Designed near real-time data processing pipeline using data orchestrations tool &techniques, supporting time-sensitive operations. Led the development of a data Lakehouse architecture, reducing data duplication by 40%.

***PROJECTS***

* Data lake and Data warehouse Implementation (04/2022-Present)
* CDR and Campaign Governance (04/2022-Present)
* Customer Data Management (04/2022-Present)
* Billing Process Digitization: Paper to Paperless (01/2020-03/2021)

***TECHNICAL SKILLS***

* **Tools & Utilities:** **AWS Services(S3, Glue, EC2, Quick Sight, Aws-Bedrock), Databricks, Azure Services.**
* **Databases: Redshift, Postgre, Ms-Sql, Oracle.**
* **Programming Languages: SQL, PySpark, Python.**
* **Data Modeling :** Star Schema, Snowflake Schema.

***PROFESIONAL EXPERIENCE***

**Data Engineer and Analytics**

**Bajaj Finserv Direct Ltd (Apr‘21 – Till)**

* Developed and implemented data lake end to end pipeline using Data Lake, S3, PySpark, Redshift with different layers for CDR and campaign response and offers data.
* Developed pipeline and ETL processes using Glue and PySpark to ingest and transform data from various sources including on-premises databases and cloud services.
* Created and implemented a data warehouse solution using Redshift to support complex reporting and analytics requirements, optimizing query performance and scalability
* Developed data quality frameworks to detect, log, and correct anomalies such as missing values, duplicates, and inconsistencies, leading to improvement in data accuracy.
* Worked closely with stakeholders to gather req and define data architecture, ensure successful project delivery. Optimized Data Warehouse performance by partitioning tables, indexing, and tuning queries.

**Data Analyst**

**Accenture India Pvt Ltd.[Quess Crop] | Mumbai (Jan’20-Apr’21)**

* Developed ETL processes using Azure Data Factory and Synapse to ingest and transform data from various sources including on-premises databases and cloud services.
* Set up data integration solutions using **SSIS** and **Azure Data Factory**, decreasing data ingestion time by 50%.
* Automated manually Daily/Weekly/Monthly task and reports extraction from email by adhering to business requirements and reducing manual efforts to 10%.
* Created **SQL** , procedures, Functions, Triggers, Views and exception handling for retrieving, manipulating, checking and migrating complex datasets in oracle to Teradata using Fast load utilize to data analysis
* Developed Ad-Hoc reports and dashboard for the Client and decision makers for Bill Strategic planning
* Provided data, reports and dashboard using Blob Storage, ADF, WinSCP, SharePoint to business team for **Email validation, SMS Validation, Paper to paperless** and Various Billing campaign.

**SQL Developer**

**Fino Payment Bank Ltd. | Mumbai (Oct’18-Jan’20)**

* Developed complex SQL queries, stored procedures, and function to support data-driven decision-making.
* Proficient in creating detailed reports and dashboards, translating complex data into actionable insights for business stakeholders
* Developed reports & dashboard weekly/monthly pan India location transactions from multiple data sources using MS-SQL,SSIS,ETL. Monitoring the payment and account related all **transaction** and solved **issues**, **Bugs**, **Error.**

**Application Support Engineer**

**Honeywell[Microsystem] | Mumbai (Apr’17-Jan’18)**

* Providing the remotely technical support to the CCTV Monitoring issues of Vodafone Store in **PAN** India any kind of issue related **DVR** over the Chat by taking **Remote Control**.
* Guide to for team Installation DVR and configuration system, HUS-NVR issues solving.

***EDUCATION & CERTIFICATION***

**Bachelor of Engineering**  **-** Mumbai University (2015)

**Certification & Training -** Microsoft Azure Fabric and AWS Bedrock